P1 1168887

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

May 13, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/455,654*
FILING DATE: *March 18, 2003*
RELATED PCT APPLICATION NUMBER: *PCT/US04/05239*

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

M. Sias

M. SIAS
Certifying Officer

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

03-19-02 60455654.031803

Please type a plus sign (+) inside this box → [+]

PTO/SB/16 (02-01)
Approved for use through 10/31/2002. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the paperwork reduction act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c)

Express Mail Label No. EV171061076US

INVENTOR(s)

Given Name (first and middle (if any))		Family Name or Surname	Residence (City and either State or Foreign Country)	
Uzi	Yehoshua	Vishkin	Rockville	MD

☐ *Additional inventors are being named on the _____ separately numbered sheets attached hereto*

TITLE OF THE INVENTION (280 characters max)

Optical Interconnect Using a Small Number of Layers

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☒ **Customer Number** 21771 → *Place Customer Number Bar Code Label here*

OR *Type Customer Number here*

☒ Firm or Individual Name	OFFICE OF TECHNOLOGY COMMERCIALIZATION				
Address	UNIVERSITY OF MARYLAND				
Address	6200 BALTIMORE AVENUE, SUITE 300				
City	RIVERDALE	State	MD	Zip	20737-1054
Country	USA	Telephone	301-403-2711	Fax	301-403-2717

ENCLOSED APPLICATION PARTS (check all that apply)

☒ Specification — Number of Pages 4
☒ Drawing(s) — Number of Sheets 6
☐ Application Data Sheet. See 37 CFR 1.76
☐ CD(s), Number
☒ Other (specify) ADVANCE NOTICE SERIAL NUMBER POSTCARD

METHOD OF PAYMENT OF FILING FEES FOR THIS PROVISIONAL APPLICATION FOR PATENT

☒ Applicant claims small entity status. See 37CFR 1.27.
☐ A check or money order is enclosed to cover the filing fees
☒ The commissioner is hereby authorized to charge filing fees and credit Deposit Account Number. 210683
☐ Payment by credit card. Form PTO-2038 is attached.
☒ The Commissioner is hereby authorized to charge any deficiency in the payment of the required fee(s) or credit any overpayment to Deposit Account No. 210683.

FILING FEE AMOUNT ($) $ 80.00

The invention was made by an agency of the United States Government or under contract with an agency of the United States Government.

☒ No.
☐ Yes, the name of the Government agency and the Government contract number are: _____

Respectfully Submitted via Express Mail,

SIGNATURE _____ **Date** March 18 2003

TYPED or PRINTED NAME James A. Poulos, III **REGISTRATION NO. (if appropriate)** 31,714

TELEPHONE 301-403-2711 **Docket Number** PS-2003-019

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is used by the public to file (and by the PTO to process) a provisional application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 8 hours to complete including gathering, preparing, aand submitting the complete provisional application to the PTO. Time will vary depending upon the needs of the individual case. Any comments on the amount of time you are required to complete this form should be sent to the Chief Information Officer, Patent and Trademark Office, Washington, D.C. 20231. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Box Provisional Application, Assistant Commissioner for Patents, Washington, D.C. 20231.

CONFIDENTIAL

Optical Interconnects Using a Small number of Layers, by Uzi Vishkin

BACKGROUND OF INVENTION Driven in part by the boom in telecommunications, optics has made tremendous strides in the 1990s. Optical interconnects inside computers are getting increasing attention as reported in N. Savage, Linking with light, IEEE Spectrum, August 2002, 32--36. Assuming that processing unit will continue to mostly be electronics-based, the closer the optical interconnect is to the processing elements, the more challenging the introduction of optics could become, due to the need to operate at high speeds and the resulting power requirements. Modern computer design put processing elements and highest level of cache memories on the same large computer chip. Such a chip needs to be manufactured using recent VLSI technology to allow for larger memories and strong interconnect to be included. The use of an optical interconnect between processing units and the first level of the cache could replace altogether the need for a large VLSI chip based on the most advanced technology. The processing elements and the caches could instead reside on several chips, and these chip could be much smaller and based on older and therefore cheaper chip technologies. If properly packaged with the optical interconnect, they could provide the same performance, but at a small fraction of the cost. For example, rather than put 64 processing+memory modules, as well as interconnect fabric, on a single .065 micron chip, we could go several generations back and use .25 micron technology for 64 (very inexpensive!) chips packaged with the optoelectronic component comprising the interconnect. To be competitive, the optoelectronic component and the overall packaging will have to be relatively inexpensive.

SUMMARY OF THE INVENTION

A new paradigm for an optical interconnect which could serve any level of the memory hierarchy, including between parallel processing elements and the first level of the cache is presented. A key attraction of optical interconnects is that optical communication channels can cross in the same plane, and that they need not be implemented using straight lines. The invention allows all the switching (or processing of data) to be done in electronics, where optics is only used to transport data. Given a plurality of modules, each comprising processing and memory elements, the interconnect provides a system of optical communication channels between every module and every other module, such that even if the optical communication channels are implemented in the plane: (i) the bending of each optical communication channel is limited, (ii) if two optical communication channels cross, their angle is not too acute (i.e., it is close to 90 degrees), (iii) only two optical communication channels can cross at the same point, (iv) the distance between any two crossing points is not too small, and (v) unless near their crossing point, the distance between two optical communication channels is not too small.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1.1 An all-to-all straight-line geometric interconnect among 32 processor+memory modules. Each module has 31 lines connecting it to each other.

Figure 1.2 is a close-up on a part of Figure 1.1. The zooming is by a factor of 2 on the X axis and the Y axis.

Figure 1.3 is a close-up by the same factor on Figure 1.2.

Figure 1.4 is a close-up by the same factor on Figure 1.3.

Figure 1.5 is a close-up by the same factor on Figure 1.4. Figure 2. The main idea which enables modifying Figure 1 into an interconnect is illustrated. The interconnect would satisfy the limited bending, not-acute angle, not-too-near crossings and the not-too near channel requirements.

Figure 2 comprises three circles. The upper circle in Figure 2 corresponds to the innermost circle in Figure 1.1. The main idea would be to modify the upper circle in Figure 2 into the lower circle in Figure 2 (and put this "patch" back into Figure 1). For methodical presentation reasons, the middle circle of Figure 2 is presented first. The middle circle illustrate how to satisfy the not-acute angle, not-too-near crossings and the not-too near channel requirements. The lower circle than illustrates how to modify further the middle circle to satisfy the limited bending requirement.

DETAILED DESCRIPTION OF THE PREFERRED EMBODYMENT

As a first approximation, Figure 1.1 depicts an all-to-all straight-line geometric interconnect among 32 processor+memory modules. Each module has 31 lines connecting it to the other modules. Figure 2 provides the main idea in order to turn Figure 1 into a useful interconnect for XMT. Suppose that:
(i) the diameter of Figure 1 was 25 centimeters (10 inches) (ii) it is implemented as a single-layer waveguide, (iii) a waveguide does not have to be a straight-line; the waveguide can be bent so that the bent part will at no point have a radius of curvature less than 50 micrometers, (iv) two waveguides can cross in the plane, preferably with a right (90 degree) angle; one alternative is to bend a waveguide over the other to avoid crossing in the same plane, (v) only two waveguides can cross at the same point and the distance between two crossing points is at least 100 micrometer, (vi) unless near their crossing point, the distance between two waveguides is never less than 100 micrometer.

Figure 2 provides a simple way to satisfy all these constraints for 32 processor+memory modules for the point at the center of Figure 1.1, where 16 lines meet. It is later noted that this way and some ad-hoc bending of lines can be used to satisfy all these constraints for 32 processor+memory modules everywhere else in Figure 1.1. This is done without

lengthening the waveguides by much. Although not detailed here, all these techniques could be extended even for 64 processor+memory modules. Figure 2 shows how to bend the 8 lines that come from the North-West quadrant so that they all run parallel to one another; the 8 lines that come from the North-East quadrant also run parallel to one another; the former 8 lines form a grid with the latter 8 lines providing all the crossings between them where no two crossings are too close. The crossings within each group of 8 lines are obtained by recursively repeating a similar grid for each group. Figure 2 depicts the crossings within the 2 groups of 8 lines, and then within the 4 groups of 4 lines and finally within the 8 groups of 2 lines. Figure 1.2-5 are provided to illustrate why the point at the center of Figure 1.1 is most problematic, and why the situation elsewhere is much easier to handle. Figure 1.2-5 were obtained by zooming on Figure 1.1. Figure 1.2 is a close-up on a part of Figure 1.1. The zooming is by a factor of 2 on the X axis and on the Y axis . The zooming in Figure 1.3 is by a factor of 4 relative to Figure 1.1 on each axis. The zooming in Figure 1.4 is by a factor of 8, and the zooming in Figure 1.5 is by a factor of 16 relative to Figure 1.1 Figure 1.5 shows that no more than 3 lines intersect at the same point. It also suggests that there is sufficient space for combining ad-hoc bending of lines with the solution of Figure 2 to satisfy all these constraints.

Depending on the exact optoelectronic technology used, the following issues which are beyond the scope of the current invention will need to be addressed: - how to get communication rates that fit the needs of the application? - the communication rates for each channel will be limited not only by the capacity of the channel but also by the capacities of the sending and receiving ends which would need to temporarily store the transmitted data; one way for regulating the aggregate rate for all channels with the same receiving end would be by using the so-called prefix-sum apparatus, as per U. Vishkin, Prefix sums and an application thereof, US Patent application 09/224,104, December 31, 1998, allowed. Each channel that needs to send data to a common destination will broadcast the size of the data, and the receiving end will issue back future time slots for the transmissions on each of the channels to ensure that the amount of data received at any point in time can be safely handled. This type of load balancing computation at the receiving end could be done by electronic hardware; our interconnect apparatus only requires that the channels use optics. - thermo-modeling: translation of optics-to-electronics and back and driving optical signals to accomplish our performance objective requires considerable power; how to evaluate the resulting heat and minimize it? - spacing: what is the correct stacking density of processor+memory modules in view of this thermo modeling? the higher the heat, the larger the diameter of the interconnect has to be to facilitate cooling; since the speed-of-light is 30cm/ns, a too large diameter could increase latencies by too much for the application. - if waveguide technology is used, what would be the most appropriate waveguide technology? Will it be silica-on-silicon? - how many crossings can we allow for each waveguide and still meet performance objectives? for a 64 module interconnect, a waveguide may cross up to 1000 others; this seems to allow a loss of no more than .05dB per crossing, which requires special attention; - how big will radiative/scattering loss be? - will the waveguide approach, or any other approach, lend itself to low-cost mass production similar to mask-based VLSI? Recall that we seek a substitute to a large on-chip design. The cost for 64 modules which are much smaller is going to be minimal, as they could rely on older VLSI technologies.

So, if the interconnect and its overall packaging become affordable, the whole approach becomes affordable as well, - will approaches other than using waveguides, such as free-space optics or fiber optics work better? The motivation of the current invention was provided by U. Vishkin, Spawn-join instruction set architecture for providing explicit multithreading (XMT) , Issued as US Patent 6,463,527 on October 8, 2002. An all-electronic interconnect was proposed in J. Nuzman and U. Vishkin, Circuit architecture for reduced-synchrony on-chip interconnect. US Provisional Patent Application Serial No. 60/0297,248, June 12, 2001.

A substantial challenge for an XMT design is to provide connectivity between the many execution units and the many cache modules on-chip. While the capacity for sending signals increases with each technology shrinkage, the latency for propagating signals down a long wire is increasing. Due to the memory model supported, memory requests can travel to any memory location on the chip. A latency cost for such memory accesses can not be avoided. Fortunately, the independence of order characteristic of XMT threading allows for such latency to be tolerated. The Nuzman-Vishkin patent application is based on supporting simultaneous requests by pipelining throughout the interconnection network. The memory subsystem interconnect (along with the hardware prefix-sum mechanism) is one of the very few global resources in an XMT design. Providing a centralized scheduling resource to coordinate communication would be very costly for a large design.

Driving a fast global clock across a deep submicron chip is also very difficult and power consumptive. The solution to both problems is to use a decentralized routing scheme. The hardware cost of tagging and local switching structures is easily supported by the benefits of such an asynchronous or loosely synchronous structure, as both the Nuzman-Vishkin invention and the current invention provide. An alternative preferred embodyment could rely on a 2-layer implementation. In this case, only two optical communication channels could cross at the same vertical point (i.e., same X coordinate and same Y coordinate, but a different Z Coordinate). Limited vertical bending of an optical communication channel in order to advance from one layer to another is allowed. The same design as in the preferred embodyment could be used, where: (i) unless near a crossing the optical communication channels are all in the same layer, and (ii) near each crossing one of the communication channels bends vertically into the other layer, and then bends back again into the first layer.

WHAT IS CLAIMED IS: ABSTRACT The invention presents a new paradigm for an all-to-all optical interconnect which could serve a variety of functions, including between parallel processing elements and the first level of the cache. The interconnect builds on the ability of optical communication channels to cross in the same plane and advance along non-straight lines.

****

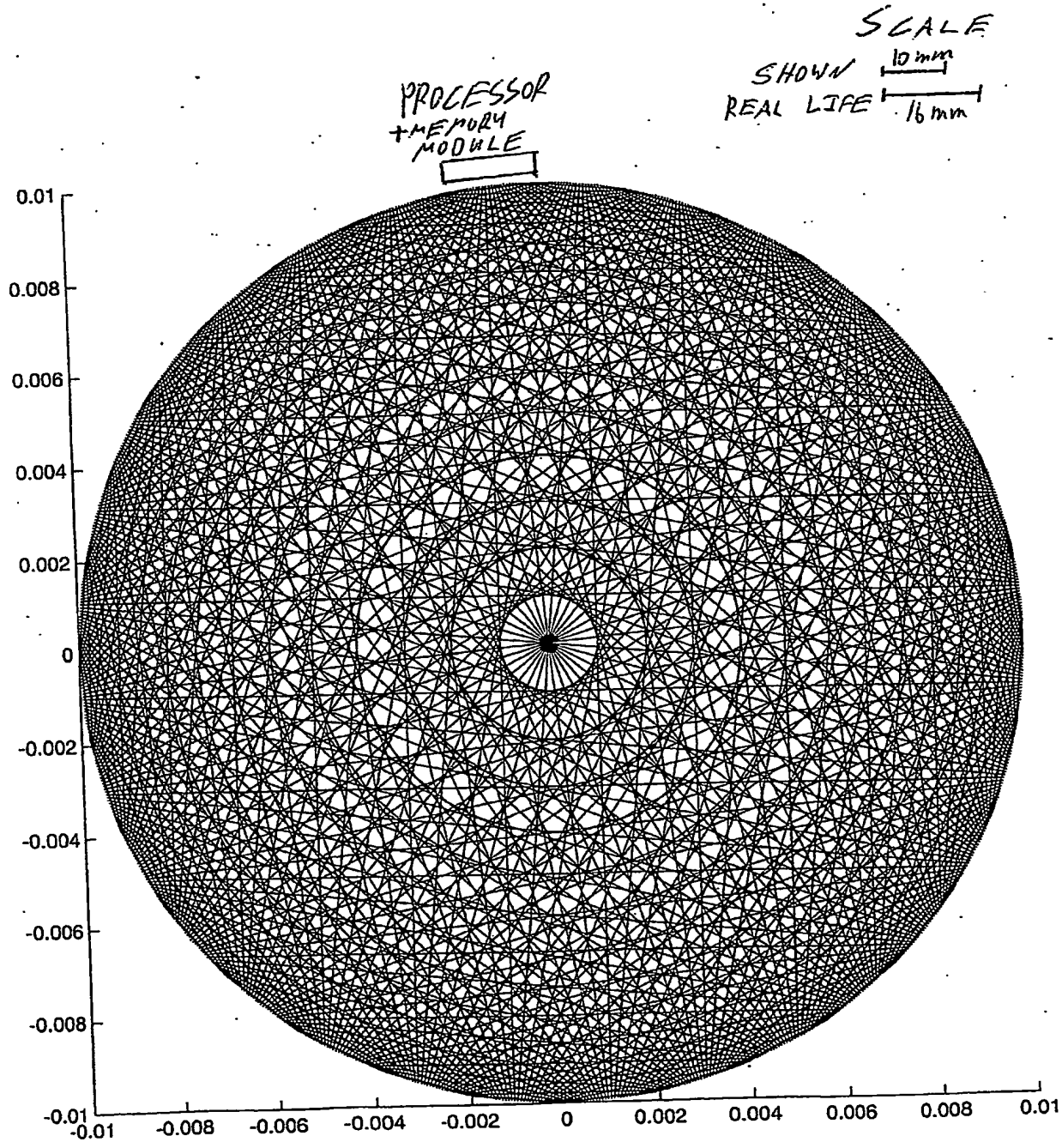
SCALE
SHOWN 10 mm
REAL LIFE 16 mm
PROCESSOR +MEMORY MODULE
0.01
0.008
0.006
0.004
0.002
0
-0.002
-0.004
-0.006
-0.008
-0.01
-0.01
-0.008
-0.006
-0.004
-0.002
0
0.002
0.004
0.006
0.008
0.01

FIG 1.1

SCALE SHOWN REAL LIFE 10 mm 3 mm

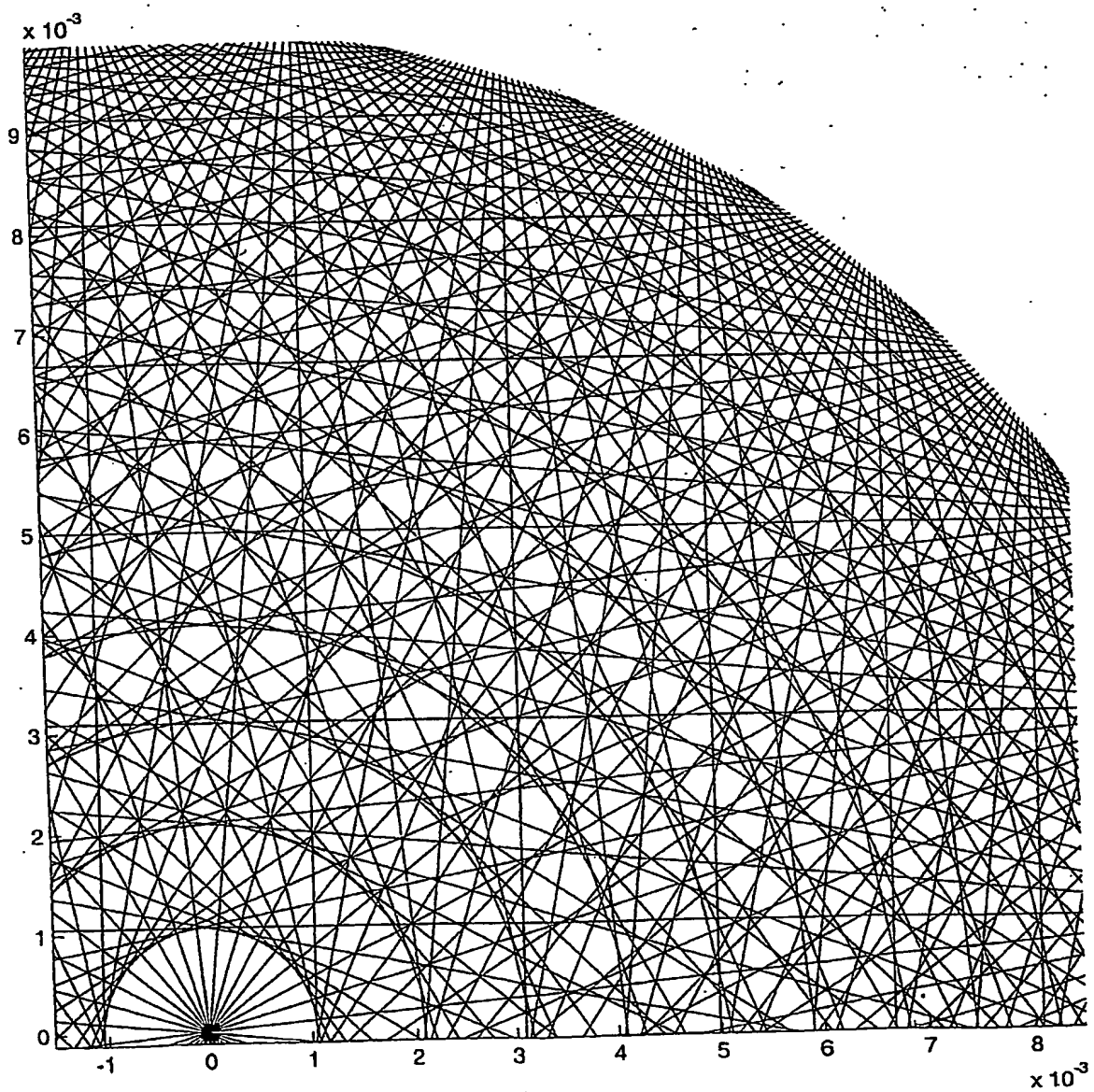

FIG 1.2

SCALE
SHOWN 10mm
REAL LIFE 4mm

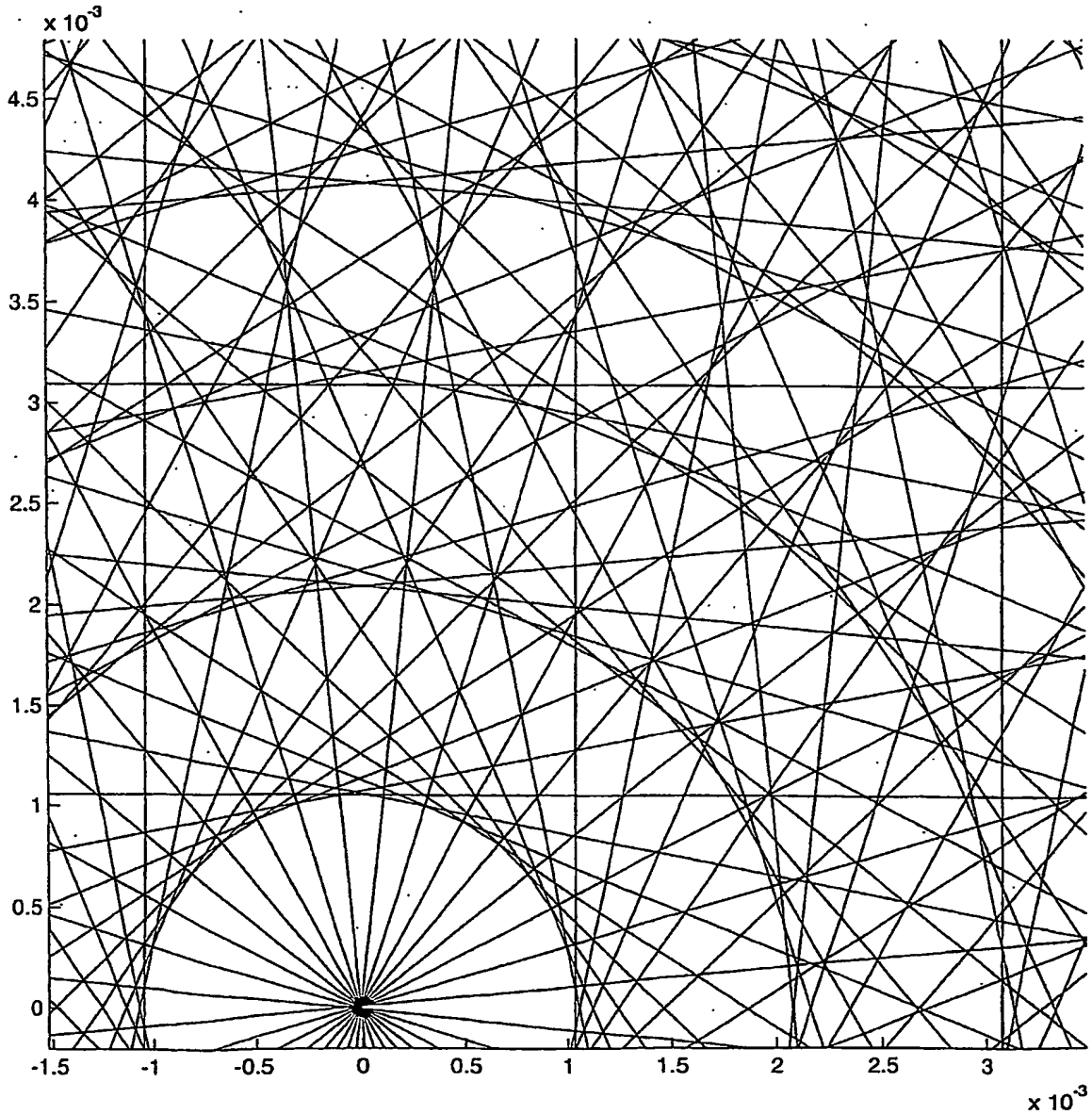

FIG 1.3

SCALE SHOWN REAL LIFE

10 mm
2 mm

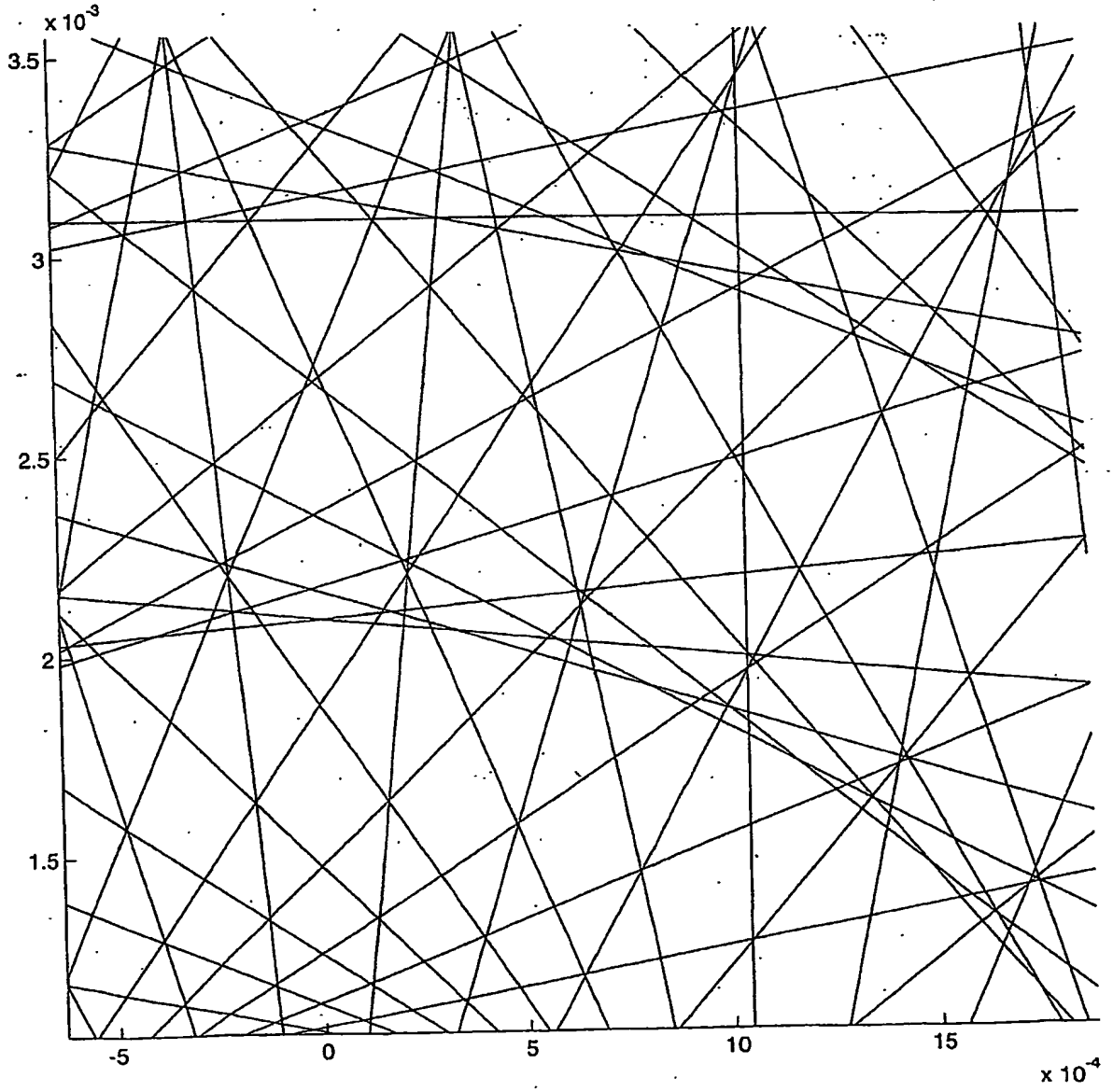

FIG 1.4

SCALE SHOWN REAL LIFE

10 mm

1 mm

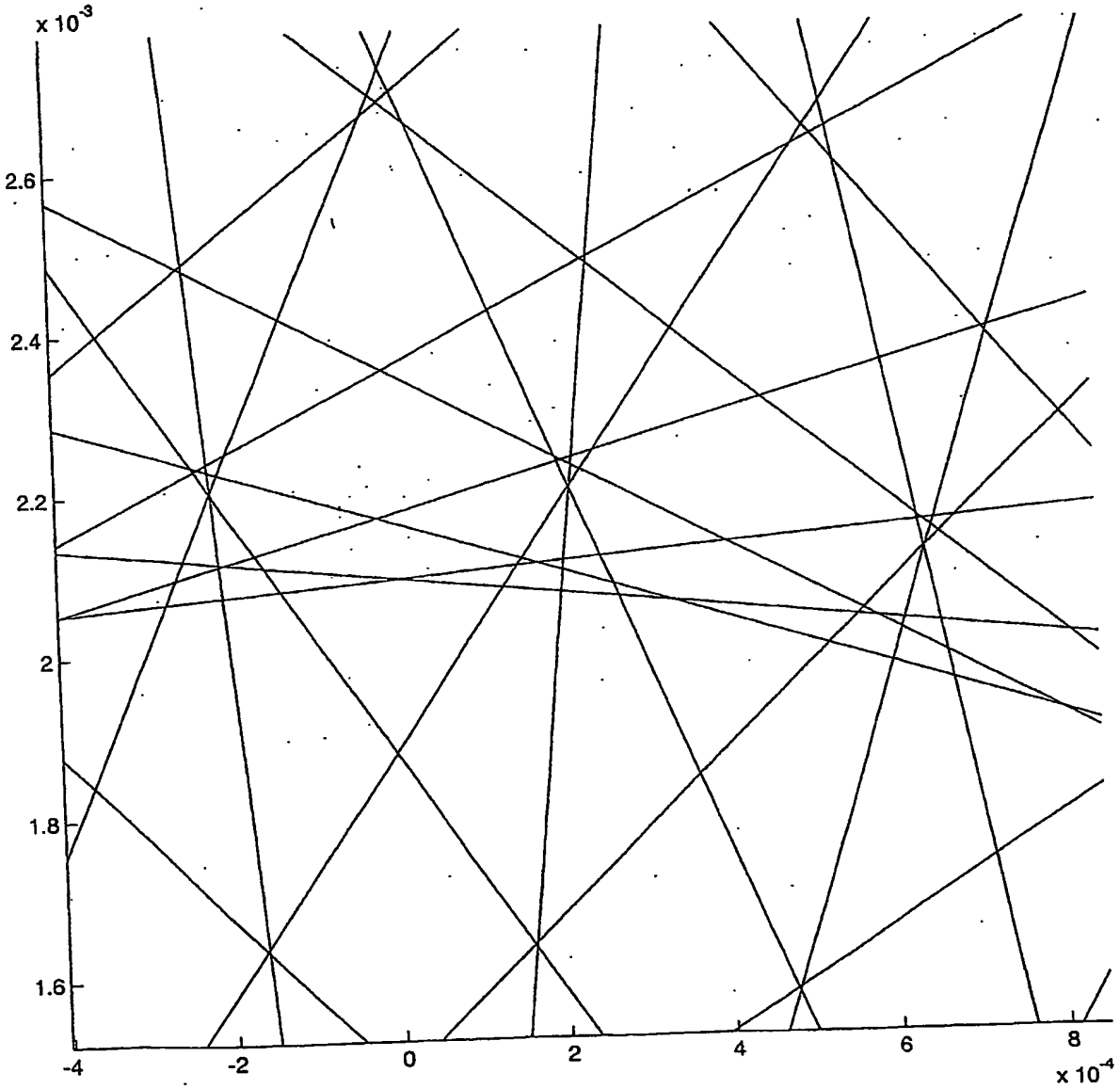

FIG 1.5

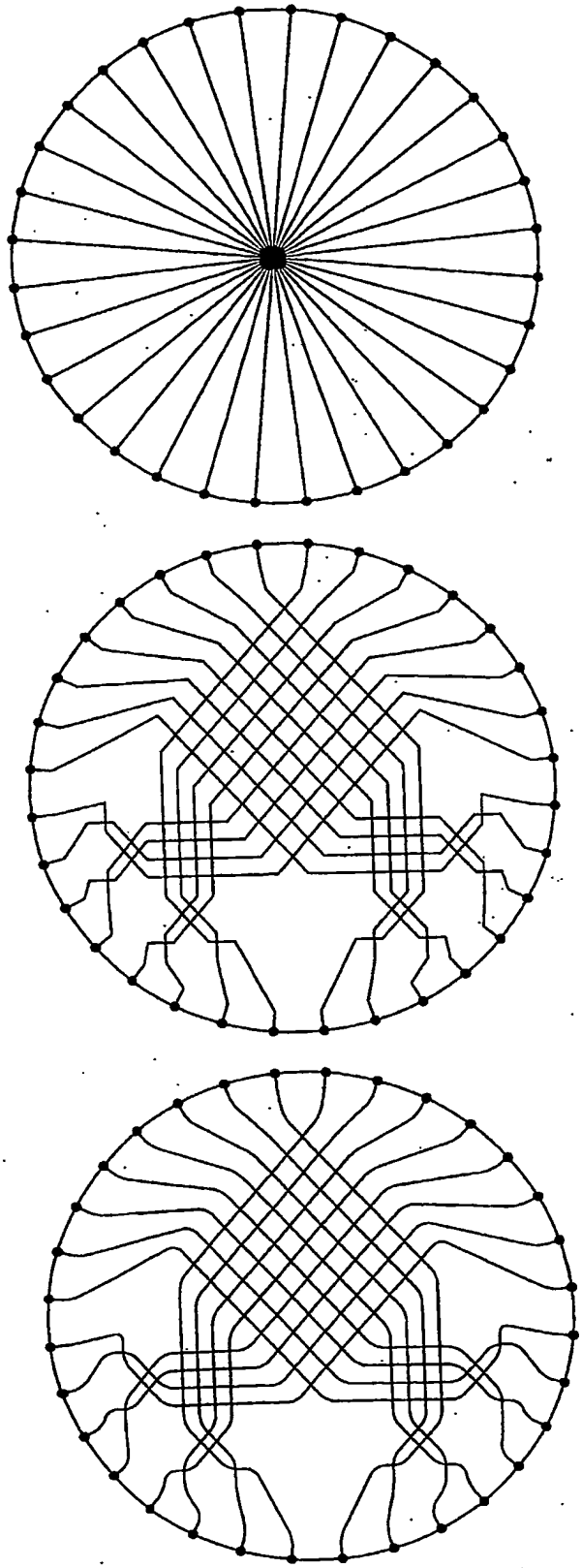

Fig. 2